00

Kongeriget Danmark

Patent application No.: PA 2000 00666

Date of filing: 19 April 2000

Applicant: (Name and address)
Statens Serum Institut
Artillerivej 5
DK-2300 København S
Denmark

Title: Nucleic acid fragments and polypeptide fragments derived from M. Tuberculosis.

IPC:-

This is to certify that the attached documents are exact copies of the above mentioned patent application as originally filed.

Patent- og Varemærkestyrelsen
Kongeriget Danmark

Patent- og Varemærkestyrelsen
Økonomi- og Erhvervsministeriet

26 November 2003

Hgs
Henrik Grye Skou

PATENT- OG VAREMÆRKESTYRELSEN

Nucleic acid fragments and polypeptide fragments derived from M. Tuberculosis

FIELD OF THE INVENTION

The present invention relates to immunologically active, polypeptide fragments derived from the *Mycobacterium tuberculosis*, vaccines and other immunologic compositions containing the fragments as immunogenic components, and methods of production and use of the polypeptides. The invention also relates to novel nucleic acid fragments derived from *M. tuberculosis* which are useful in the preparation of the polypeptide fragments of the invention or in the diagnosis of infection with *M. tuberculosis*.

Background of the invention

Human tuberculosis (hereinafter designated "TB") caused by *Mycobacterium tuberculosis* is a severe global health problem responsible for approximately 3 million deaths annually, according to the WHO. The worldwide incidence of new TB cases has been progressively falling for the last decade but the recent years has markedly changed this trend due to the advent of AIDS and the appearance of multidrug resistant strains of *M. tuberculosis*.

The only vaccine presently available for clinical use is BCG, a vaccine which efficacy remains a matter of controversy. BCG generally induces a high level of acquired resistance in animal models of TB, but several human trials in developing countries have failed to demonstrate significant protection. Notably, BCG is not approved by the FDA for use in the United States. Although the BCG vaccine is widely used, some countries including the USA never introduced it for use in general population vaccination programmes, one reason being that vaccination with BCG interferes with the use of tuberculin skin testings for diagnosing tuberculosis and for use in population surveys.

This makes the development of a new and improved vaccine against TB an urgent matter which has been given a very high priority by the WHO.

It is an object of the invention to provide novel antigens which are effective as components in a subunit vaccine against TB or which are useful as components in diagnostic compositions for the detection of infection with mycobacteria, especially

virulence-associated mycobacteria. The novel antigens may also be important drug targets.

Summary of the invention

Name of antigen	Nucleotide sequence SEQ ID NO:	Amino acid sequence SEQ ID NO:
Rv0284	1	2
Rv0284 3' part / c-terminal	3	4
Rv0285	5	6
Rv3878 (ORF11)	7	8

Detailed disclosure of the invention

The present invention is i.a. based on the identification and characterisation of a number of previously uncharacterised antigens from M. tuberculosis as presented in the examples. It is well-known in the art that T-cell epitopes are responsible for the elicitation of the acquired immunity against TB. Since such T-cell epitopes are linear and are known to have a minimum length of 6 amino acid residues, the present invention is especially concerned with the identification and utilisation of such T-cell epitopes.

As illustrated in example 2, Rv0284 (SEQ ID NO: 2) causes a marked release of IFN-γ from PBMC withdrawn from TB patients from half of the donors, and in a majority of PPD positive healthy donors without any increase in PPD negative healthy donors. Rv0284 further stimulates T cell lines from PPD positive donors to release IFN-γ and induces a DTH reaction in guinea pigs aerosol infected with M. tuberculosis. This indicates that Rv0284 is highly biologically active and recognised by PPD positive donors and TB patients.

Thus, one aspect of the invention relates to a substantially pure polypeptide fragment which comprises an amino acid sequence as shown in SEQ ID NO: 2 or comprises an amino acid sequence analogue having a sequence identity with the polypeptide fragment shown in SEQ ID NO: 2 of at least 70% and at the same time being immunologically equivalent to the polypeptide fragment shown in SEQ ID NO: 2

A related aspect of the invention relates to a substantially pure polypeptide fragment which comprises a T-cell epitope of the amino acid sequence as shown in SEQ ID NO: 2 and at the same time being immunologically equivalent to the polypeptide shown in SEQ ID NO: 2.

As illustrated in example 2, Rv0285 (SEQ ID NO: 6) stimulates T cell lines from PPD positive donors to release IFN-γ to a level of close to the release caused by PPD. . Rv0285 further induces a DTH reaction in guinea pigs aerosol infected with M. tuberculosis. This indicates that Rv0285 is highly biologically active and recognised by PPD positive donors and TB patients.

As illustrated in example 2, Rv3878 (SEQ ID NO: 8) causes a marked release of IFN-γ from PBMC withdrawn from TB patients and in PPD positive healthy donors without any increase in PPD negative healthy donors. Rv3878 further stimulates T cell lines from PPD positive donors to release IFN-γ to a level resembling the release caused by PPD and Rv0285 also induces a DTH reaction in guinea pigs aerosol infected with M. tuberculosis.. This indicates that Rv3878 is highly biologically active and recognised by PPD positive donors and TB patients.

A polypeptide fragment is considered to be "immunologically equivalent" to a polypeptide disclosed in the present invention, if it

1) induces *in vitro* recall response determined by release of IFN-γ of at least 30% of the release induced by the polypeptide disclosed from Peripheral Blood Mononuclear Cells (PBMC) or whole blood withdrawn from TB patients 0-6 months after diagnosis, or PPD positive individual, the inductions being performed by the addition of the polypeptide disclosed and the polypeptide fragment to two individual suspensions comprising about 1.0 to 2.5 x 10^5 PBMC or whole blood cells, the addition of the polypeptides resulting in a concentration of not more than 20 μg per ml suspension, the release of IFN-γ being assessable by determination of IFN-γ in supernatant harvested 5 days after the addition of the polypeptide to the suspension; or

2) it induces a positive DTH response determined by intradermal injections or local application patchs of at most 100 μg of the polypeptide disclosed and of the polypeptide fragment to an individual who is clinically or subclinically infected with a virulent *Mycobacterium*, and the polypeptide fragment causing a response diameter that is at least

50% of that caused by the polypeptide disclosed measured 72-96 hours after the injections or applications.

Each polypeptide disclosed in the present application is characterised by specific amino acid and nucleic acid sequences. It will be understood that such sequences include analogues and variants produced by recombinant methods wherein such nucleic acid and polypeptide sequences have been modified by substitution, insertion, addition and/or deletion of one or more nucleotides in said nucleic acid sequences to cause the substitution, insertion, addition or deletion of one or more amino acid residues in the recombinant polypeptide.

In both immunodiagnostics and vaccine preparation, it is often possible and practical to prepare antigens from segments of a known immunogenic protein or polypeptide. Certain epitopic regions may be used to produce responses similar to those produced by the entire antigenic polypeptide. Potential antigenic or immunogenic regions may be identified by any of a number of approaches, e.g., Jameson-Wolf or Kyte-Doolittle antigenicity analyses or Hopp and Woods (1981) hydrophobicity analysis (see, e.g., Jameson and Wolf, 1988; Kyte and Doolittle, 1982; or U.S. Patent No. 4,554,101). Hydrophobicity analysis assigns average hydrophilicity values to each amino acid residue from these values average hydrophilicities can be calculated and regions of greatest hydrophilicity determined. Using one or more of these methods, regions of predicted antigenicity may be derived from the amino acid sequence assigned to the polypeptides of the invention.

Alternatively, in order to identify relevant T-cell epitopes which are recognized during an immune response, it is also possible to use a "brute force" method: Since T-cell epitopes are linear, deletion mutants of polypeptides having SEQ ID NO: 2, 4, 6, or 8 will, if constructed systematically, reveal what regions of the polypeptides are essential in immune recognition, e.g. by subjecting these deletion mutants to the IFN-γ assay described herein. Another method utilises overlapping oligomers (preferably synthetic having a length of e.g. 20 amino acid residues) derived from polypeptides having SEQ ID NO: 2, 4, 6, or 8. Some of these will give a positive response in the IFN-γ assay whereas others will not.

In a preferred embodiment of the invention, the polypeptide fragment of the invention comprises an epitope for a T-helper cell.

Although the minimum length of a T-cell epitope has been shown to be at least 6 amino acids, it is normal that such epitopes are constituted of longer stretches of amino acids. Hence it is preferred that the polypeptide fragment of the invention has a length of at least 7 amino acid residues, such as at least 8, at least 9, at least 10, at least 12, at least 14, at least 16, at least 18, at least 20, at least 22, at least 24, and at least 30 amino acid residues.

By producing fusion polypeptides, superior characteristics of the polypeptide fragments of the invention can be achieved. For instance, fusion partners which facilitate export of the polypeptide when produced recombinantly, fusion partners which facilitate purification of the polypeptide, and fusion partners which enhance the immunogenicity of the polypeptide fragment of the invention are all interesting possibilities. Therefore, the invention also pertains to a fusion polypeptide comprising at least one polypeptide fragment defined above and at least one fusion partner. The fusion partner can, in order to enhance immunogenicity, e.g. be selected from the group consisting of another polypeptide fragment as defined above (so as to allow for multiple expression of relevant epitopes), and an other polypeptide derived from a bacterium belonging to the tuberculosis complex, such as ESAT-6, MPB64, MPT64, and MPB59 or at least one T-cell epitope of any of these antigens. Other immunogenicity enhancing polypeptides which could serve as fusion partners are T-cell epitopes (e.g. derived from the polypeptides ESAT-6, MPB64, MPT64, or MPB59) or other immunogenic epitopes enhancing the immunogenicity of the target gene product, e.g. lymphokines such as INF-γ, IL-2 and IL-12. In order to facilitate expression and/or purification the fusion partner can e.g. be a bacterial fimbrial protein, e.g. the pilus components pilin and papA; protein A; the ZZ-peptide (ZZ-fusions are marketed by Pharmacia in Sweden); the maltose binding protein; gluthatione S-transferase; β-galactosidase; or poly-histidine.

Other interesting fusion partners are polypeptides which are lipidated and thereby effect that the immunogenic polypeptide is presented in a suitable manner to the immune system. This effect is e.g. known from vaccines based on the Borrelia burgdorferi OspA polypeptide, wherein the lipidated membrane anchor in the polypeptide confers a self-adjuvating effect to the polypeptide (which is natively lipidated) when isolated from cells producing it. In contrast, the OspA polypeptide is relatively silent immunologically when prepared without the lipidation anchor.

A substantially pure polypeptide according to any of the preceding clams for use as a pharmaceutical.

Use of a substantially pure polypeptide according to any of the preceding claims for the preparation of a pharmaceutical composition for the diagnosis of tuberculosis caused by *Mycobacterium tuberculosis*, *Mycobacterium africanum* or *Mycobacterium bovis*.

Use of a substantially pure polypeptide according to any of the preceding claims for the preparation of a pharmaceutical composition for the vaccination against infections caused by *Mycobacterium tuberculosis*, *Mycobacterium africanum* or *Mycobacterium bovis*.

Another part of the invention pertains to an immunologic composition comprising a polypeptide or fusion polypeptide according to the invention. In order to ensure optimum performance of such an immunologic composition it is preferred that it comprises an immunologically and pharmaceutically acceptable carrier, vehicle or adjuvant.

Suitable carriers are selected from the group consisting of a polymer to which the polypeptide(s) is/are bound by hydrophobic non-covalent interaction, such as a plastic, e.g. polystyrene, or a polymer to which the polypeptide(s) is/are covalently bound, such as a polysaccharide, or a polypeptide, e.g. bovine serum albumin, or keyhole limpet haemocyanin. Suitable vehicles are selected from the group consisting of a diluent and a suspending agent. The adjuvant is preferably selected from the group consisting of dimethyldioctadecylammonium bromide (DDA), Quil A, poly I:C, Freund's incomplete adjuvant, IFN-γ, IL-2, IL-12, monophosphoryl lipid A (MPL), and muramyl dipeptide (MDP).

A preferred immunologic composition according to the present invention comprising at least two different polypeptide fragments, each different polypeptide fragment being a polypeptide or a fusion polypeptide defined above. It is preferred that the immunologic composition comprises between 3-20 different polypeptide fragments or fusion polypeptides.

Such an immunologic composition may preferably be in the form of a vaccine or in the form of a skin test reagent.

In line with the above, the invention therefore also pertain to a method for producing an immunologic composition according to the invention, the method comprising preparing, synthesizing or isolating a polypeptide according to the invention, and solubilizing or dispersing the polypeptide in a medium for a vaccine, and optionally adding other M. tuberculosis antigens and/or a carrier, vehicle and/or adjuvant substance.

Each of the polypeptides may be characterised by specific amino acid and nucleic acid sequences. It will be understood that such sequences include analogues and variants produced by recombinant methods wherein such nucleic acid and polypeptide sequences have been modified by substitution, insertion, addition and/or deletion of one or more nucleotides in said nucleic acid sequences to cause the substitution, insertion, addition or deletion of one or more amino acid residues in the recombinant polypeptide. A preferred nucleotide sequence encoding a polypeptide of the invention is a nucleotide sequence which

1) is a DNA sequence selected from the group consisting of SEQ ID NOs: 1, 3, 5, and 7 or an analogue of said sequence which hybridises with any DNA sequence complementary to DNA sequences shown in SEQ ID NOs: 1, 3, 5, or 7 or a specific part thereof, preferably under stringent hybridisation conditions. By stringent conditions is understood, as defined in the art, 5-10°C under the melting point T_m, cf. Sambrook et al, 1989, pages 11.45-11.49, and/or

2) encodes a polypeptide, the amino acid sequence of which has a 80% sequence identity with an amino acid sequence selected from the group consisting of SEQ ID NOs: 2, 4, 6, and 8 and/or

3) constitutes a subsequence of any of the above mentioned DNA sequences, and/or

4) constitutes a subsequence of any of the above mentioned polypeptide sequences.

The terms "analogue" or "subsequence" when used in connection with the DNA fragments of the invention are thus intended to indicate a nucleotide sequence which encodes a polypeptide exhibiting identical or substantially identical immunological properties to a polypeptide encoded by the DNA fragment of the invention shown in any of SEQ ID NOs: 1, 3, 5, or 7, allowing for minor variations which do not have an adverse effect on the

ligand binding properties and/or biological function and/or immunogenicity as compared to any of the polypeptides of the invention or which give interesting and useful novel binding properties or biological functions and immunogenicities etc. of the analogue and/or subsequence. The analogous DNA fragment or DNA sequence may be derived from a bacterium, a mammal, or a human or may be partially or completely of synthetic origin. The analogue and/or subsequence may also be derived through the use of recombinant DNA techniques.

Furthermore, the terms "analogue" and "subsequence" are intended to allow for variations in the sequence such as substitution, insertion (including introns), addition, deletion and rearrangement of one or more nucleotides, which variations do not have any substantial effect on the polypeptide encoded by a DNA fragment or a subsequence thereof. The term "substitution" is intended to mean the replacement of one or more nucleotides in the full nucleotide sequence with one or more different nucleotides, "addition" is understood to mean the addition of one or more nucleotides at either end of the full nucleotide sequence, "insertion" is intended to mean the introduction of one or more nucleotides within the full nucleotide sequence, "deletion" is intended to indicate that one or more nucleotides have been deleted from the full nucleotide sequence whether at either end of the sequence or at any suitable point within it, and "rearrangement" is intended to mean that two or more nucleotide residues have been exchanged with each other.

It is well known that the same amino acid may be encoded by various codons, the codon usage being related, *inter alia*, to the preference of the organisms in question expressing the nucleotide sequence. Thus, at least one nucleotide or codon of a DNA fragment of the invention may be exchanged by others which, when expressed, result in a polypeptide identical or substantially identical to the polypeptide encoded by the DNA fragment in question.

When the term nucleic acid is used in the following, it should be understood that for the number of purposes where nucleic acid can be substituted with DNA or RNA, which will be apparent for the person skilled in the art. For the purposes of hybridization, PNA or LNA may be used instead of DNA or nucleic acid. As DNA is the most frequently used transfection material DNA is the preferred nucleic acid.

The invention also relates to a replicable expression vector which comprises a nucleic acid fragment defined above, especially a vector which comprises a nucleic acid fragment encoding a polypeptide fragment of the invention.

The vector may be any vector which may conveniently be subjected to recombinant DNA procedures, and the choice of vector will often depend on the host cell into which it is to be introduced. Thus, the vector may be an autonomously replicating vector, i.e. a vector which exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication; examples of such a vector are a plasmid, phage, cosmid, mini-chromosome or virus. Alternatively, the vector may be one which, when introduced in a host cell, is integrated in the host cell genome and replicated together with the chromosome(s) into which it has been integrated.

Expression vectors may be constructed to include any of the DNA segments disclosed herein. Such DNA might encode an antigenic protein specific for virulent strains of mycobacteria or even hybridization probes for detecting mycobacteria nucleic acids in samples. Longer or shorter DNA segments could be used, depending on the antigenic protein desired. Epitopic regions of the proteins expressed or encoded by the disclosed DNA could be included as relatively short segments of DNA. A wide variety of expression vectors is possible including, for example, DNA segments encoding reporter gene products useful for identification of heterologous gene products and/or resistance genes such as antibiotic resistance genes which may be useful in identifying transformed cells.

The vector of the invention may be used to transform cells so as to allow propagation of the nucleic acid fragments of the invention or so as to allow expression of the polypeptide fragments of the invention. Hence, the invention also pertains to a transformed cell harbouring at least one such vector according to the invention, said cell being one which does not natively harbour the vector and/or the nucleic acid fragment of the invention contained therein. Such a transformed cell (which is also a part of the invention) may be any suitable bacterial host cell or any other type of cell such as a unicellular eukaryotic organism, a fungus or yeast, or a cell derived from a multicellular organism, e.g. an animal or a plant. It is especially in cases where glycosylation is desired that a mammalian cell is used, although glycosylation of proteins is a rare event in prokaryotes. Normally, however, a prokaryotic cell is preferred such as a bacterium belonging to the genera Mycobacterium, Salmonella, Pseudomonas, Bacillus and Eschericia. It is preferred that the

transformed cell is an E. coli, B. subtilis, or M. bovis BCG cell, and it is especially preferred that the transformed cell expresses a polypeptide according of the invention. The latter opens for the possibility to produce the polypeptide of the invention by simply recovering it from the culture containing the transformed cell. In the most preferred embodiment of this part of the invention the transformed cell is Mycobacterium bovis BCG strain: Danish 1331, which is the Mycobacterium bovis strain Copenhagen from the Copenhagen BCG Laboratory, Statens Seruminstitut, Denmark.

Examples

Example 1: Cloning and expression of Rv0284, Rv0285 and Rv3878

The coding region of Rv0285, Rv3878 and the 3'-part (380 bp) of Rv0284 were amplified by PCR using following primer sets:

Rv0284-F: CTG AGA TCT CAG GTA CCG GAT TCG CCG
BglII

Rv0284-R: CTC CCA TGG TCA TGA CTG ACT CCC CTT
NcoI

Rv0285-F: CTG AGA TCT ATG ACG TTG CGA GTG GTT
BglII

Rv0285-R: CTC CCA TGG TCA GCC GCC CAC GAC CCC
NcoI

Rv3878-F: CTG AGA TCT GCT ACT GTT AAC AGA TCG
BglII

Rv3878-R: CCG CTC GAG CTA CAA CGT TGT GGT TGT
XhoI

PCR reactions contained 10 ng of M. tuberculosis H37Rv DNA in 1x low salt Taq$^+$ buffer from Stratagene supplemented with 250 mM of each of the four nucleotides (Boehringer Mannheim), 0.5 mg/ml BSA (IgG technology), 1% DMSO (Merck), 5 pmoles of each primer and 0.5 unit Taq$^+$ DNA polymerase (Stratagene) in 10 µl reaction volume. Reactions were initially heated to 94°C for 15 sec, followed by 30 cycles of 94°C for 30 sec, 55°C for 30 sec and 72°C for 90 sec, and finally by 72°C for 5 min.

The PCR fragments were cloned into the TA cloning vector pCR2.1 (Invitrogen) and then transferred to the pMCT3 expression vector at the restriction sites indicated by the primers above. The resulted recombinant antigens carried 6-histidine residues at the N-terminal. All clones were confirmed by DNA sequencing.

To express his-tagged recombinant antigens, 100 ml of an overnight culture of XL-1 blue carrying the pMCT3 construct was added to 900 ml of LB-media containing 100 µg/ml

ampicillin, grown at 37°C with shaking. 1 mM IPTG was added at OD600 =0.4-0.6 and the culture was incubated for additional 3 - 16 hours before harvesting of cells.

For purification, the cell pellet was resuspended in 20 ml of Sonication buffer (20 mM Tris-Cl, pH 8.0, 0.5 M NaCl, 10% Glycerol, 5 mM β-ME, 0.01% Tween 20 and 1 mM imidazole). Cells were lysed and DNA was digested by treating with lysozyme (0.1 mg/ml) and DNase I (2.5 µg/ml) at room temperature for 20 min with gentle agitation. The recombinant protein was bring to solution by adding 80 ml of Sonication Buffer containing 8 M urea and sonicated the sample 5 x 30 sec, with 30 sec pausing between the pulses.

After centrifugation, the lysate was applied to a 5 ml TALON column (Clonetech). The column was then washed with 25 ml of urea containing Sonication buffer, and the bound protein was eluted by imidazole steps (5, 10, 20, 40 and 100 mM) in the same buffer. The fractions were analyzed by silver stained SDS-PAGE, and recombinant protein containing fractions were pooled and dialyzed against 3 x 1 L of 10 mM Tris-Cl (pH 8.0), 0.15 M NaCl and 0.1% SDS. Two mg of TALON purified recombinant antigen was subjected to SDS-PAGE on an 16 x 16 cm gel. After separation, the recombinant antigen band was cut out and the protein was eluted by an Model 422 Electro-Eluter (Bio-Rad). SDS was removed from eluted protein by Chloroform/Methanol extraction.

Example 2: Biological activity of the recombinant antigens.

The purified recombinant proteins were screened for the ability to induce a T cell response measured as IFN-γ release. A preliminary screening involved testing of the IFN-γ induction of T cell lines generated from PPD positive donors. This test was followed by measuring the response in PBMC preparations obtained from TB patients, PPD positive as well as negative healthy donors.

Interferon-γ induction of T cell lines

Human donors: PBMC were obtained from healthy donors with a positive *in vitro* response to PPD.

T cell line preparation: T cell lines were prepared by culturing 5 x 10^6 freshly isolated PBMC/ml with viable *M. tuberculosis* at a ratio of 5 bacteria per macrophage in a total volume of 1 ml. The cells were cultured in RPMI 1640 medium (Gibco, Grand Island, N.Y) supplemented with HEPES, and 10% heat-inactivated NHS. After 7 days in culture at 37 °C and 5% CO_2, T cells were supplemented with 50 U/ml of r-IL-2 (Boehringer Mannheim)

for approximately 7 days. Finally, the T cell lines were tested for reactivity against the recombinant antigens by stimulating 1-5 x 10^5 cells/ml with 5 µg/ml of PPD, 3 µg/ml of rRv0284ct, 5 µg/ml of rRv0285, or 2.5 µg/ml of rRv3878 in the presence of 5 x 10^5 autologous antigen-presenting cells/ml. No antigen (No ag) and PHA were used as negative and positive controls, respectively. The supernatants were harvested after 4 days of culture and stored at -80 °C until the presence of IFN-γ were analysed.

Cytokine analysis: Interferon-γ (IFN-γ) was detected with a standard sandwich ELISA technique using a commercially available pair of monoclonal antibodies (Endogen, MA, US) and used according to the manufacturer's instruction. Recombinant IFN-γ (Endogen, MA, US) was used as a standard. All data are means of duplicate wells and the variation between the wells did not exceed 10 % of the mean. Responses obtained with two different T cell lines are shown in Table 1.

As shown in Table 1, high levels of IFN-γ release are observed after stimulation with the recombinant antigens ranging from 33% (rRv0284ct) to 83% (rRv3878) of the response seen after stimulation with PPD.

Table 1. Stimulation of two T cell lines with recombinant rRv0284ct, rRv0285, and rRv3878. Responses to PHA and PPD are shown for comparison. Results are presented as pg IFN-γ/ml.

T cell line

Donor	No ag	PHA (1 µg/ml)	PPD (5 µg/ml)	rRv0284ct (3 µg/ml)	rRv0285 (5 µg/ml)	rRv3878 (2.5 µg/ml)
1	50	2975	2742	914	2019	1072
2	50	1482	803	352	548	667

Interferon-γ release from PBMC isolated from human TB patients and PPD positive and negative healthy donors

Human donors: PBMC were obtained from healthy donors with a positive *in vitro* response to purified protein derivative (PPD) or healthy donors with a negative *in vitro* response to PPD. PBMC were also obtained from TB patients with microscopy or culture proven infection. Blood samples were drawn from TB patients 0-6 months after diagnosis.

Lymphocyte preparations and cell culture: PBMC were freshly isolated by gradient centrifugation of heparinized blood on Lymphoprep (Nycomed, Oslo, Norway) and stored

in liquid nitrogen until use. The cells were resuspended in complete RPMI 1640 medium (Gibco BRL, Life Technologies) supplemented with 1% penicillin/streptomycin (Gibco BRL, Life Technologies), 1% non-essential-amino acids (FLOW, ICN Biomedicals, CA, USA), and 10% heat-inactivated normal human AB serum (NHS). The viability and number of the cells were determined by Nigrosin staining. Cell cultures were established with 1.25×10^5 PBMCs in 100 µl in microtitre plates (Nunc, Roskilde, Denmark) and stimulated with 5 µg/ml PPD or rRv0284ct and rRv3878 in a final concentration of 2.5 and 5 µg/ml, respectively. No ag was used as a negative control, whereas phytohaemagglutinin (PHA) was used as a positive control. Moreover, the response to a well-known T cell antigen, ESAT-6, was included for comparison. Supernatants for the analysis of secreted cytokines were harvested after 5 days of culture, pooled, and stored at -80 °C until use.

Cytokine analysis: IFN-γ was detected as above. Responses obtained with PBMCs from 14 individual donors are shown in Table 2.

As shown in Table 2, stimulation of PBMC from TB patients as well as PPD positive donors with both rRv0284ct and rRv3878 resulted in a marked release of IFN-γ with 55% of the donors recognising the recombinant antigens at a level of more than 500 pg/ml. As expected, none of the recombinant antigens gave rise to IFN-γ release in PPD negative donors.

Table 2. Stimulation of PBMCs from 4 TB patients, 7 PPD positive healthy donors, and 3 PPD negative healthy donors with recombinant antigen. Responses to PHA, PPD, and ESAT-6 are shown for comparison. Results are given as pg IFN-γ/ml.

TB patients

Donor	No ag	PHA (1 µg/ml)	PPD (5 µg/ml)	ESAT-6 (5 µg/ml)	rRv0284ct (2.5 µg/ml)	rRv3878 (5 µg/ml)
1	3	4541	4074	2154	809	3
2	92	3408	4891	611	236	2029
3	5	5282	4647	2827	308	149
4	10	4531	2077	38	140	287

PPD positive healthy donors

Donor	No ag	PHA (1 µg/ml)	PPD (5 µg/ml)	ESAT-6 (5 µg/ml)	rRv0284ct (2.5 µg/ml)	rRv3878 (5 µg/ml)
1	74	5413	3339	0	382	77
2	14	5614	3852	198	1324	633
3	7	6165	5808	4	2951	2732
4	63	6532	6314	1567	3009	3482
5	43	4733	6195	1272	5166	2589
6	5	3809	2582	15	5	71
7	31	6716	2275	424	1449	832

PPD negative healthy donors

Donor	No ag	PHA (1 µg/ml)	PPD (5 µg/ml)	ESAT-6 (5 µg/ml)	rRv0284ct (2.5 µg/ml)	rRV3878 (5 µg/ml)
1	0	3354	113	0	269	17
2	0	3803	563	0	22	0
3	0	3446	525	10	203	34

Together these analyses using T cell lines and PBMC, respectively, indicate that rRv0284ct, rRv0285, and rRv3878 are highly biologically active and recognised by PPD positive donors and TB patients.

Skin test reaction in TB infected guinea pigs

The skin test reactivity of the recombinant antigens was tested in *M. tuberculosis* infected guinea pigs. A group of 5 female outbreed guinea pigs of the Dunkin Hartley strain (Møllegaard Breeding and Research Center A/S, Lille Skensved, Denmark) were infected by the aerosol route in an exposure chamber of a Glas-Col® Inhalation Exposure System, which was calibrated to deliver approximately 20-25 *M. tuberculosis* Erdman bacilli into the lungs of each animal. As a control, the skin test reactivity of uninfected guinea pigs

was tested. Skin tests were performed 28 days after infection with injection of 5 μg of rRv0284ct, rRv0285, or rRv3878. As a positive control, the guinea pigs were sensitised with 10 tuberculin units (TU) of PPD (1TU = 0.02 μg) whereas injection of phosphate-buffered saline (PBS) was used as a negative control. Skin test responses (diameter of erythema) were read 24 h later by two experienced examinations and the results were expressed as the mean of the two readings. The variation between the two readings was less than 10%. Skin test responses larger than 5 mm were regarded as positive.

As seen in Table 3, injection of rRv3878 induced a marked Delayed Type Hypersensitivity (DTH) reaction at the same level as after injection with PPD. rRv0284ct and rRv0285 resulted in a highly significant DTH reaction ($P < 0.005$; Tukey test). As expected, none of the antigens induced non-specific response in uninfected guinea pigs (Table 4).

Table 3. DTH erythema diameter (shown in mm) in guinea pigs aerosol infected with *M. tuberculosis* after stimulation with recombinant antigens.

Antigen[a]	Skin reaction (mm)[b]	SEM
PBS	3.10	0.30
PPD	13.10	1.18
rRv0284ct	8.40	0.45
rRv0285	7.00	1.08
rRv3878	14.56	1.05

[a] The recombinant antigens were tested in a concentration of 5 μg, whereas 10 TU of PPD were used.
[b] The skin reactions are measured in mm erythema 24 h after intradermal injection. The values are the mean of erythema diameter of five animals and the SEM are indicated. The values for rRv3878 are the mean of four animals.

Table 4. DTH erythema diameter (shown in mm) in non-infected guinea pigs after stimulation with recombinant antigens.

Antigen[a]	Skin reaction (mm)[b]	SEM
PBS	2.60	0.36
PPD	3.00	0.44
rRv0284ct	2.5	0.18
rRv0285	3.45	0.74
rRv3878	2.5	0.18

[a] The recombinant antigens were tested in a concentration of 5 μg, whereas 10 TU of PPD were used.
[b] The skin reactions are measured in mm erythema 24 h after intradermal injection. The values are the mean of erythema diameter of five animals and the SEM are indicated.

Claims

1. A substantially pure polypeptide fragment which comprises an amino acid sequence as shown in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 or SEQ ID NO: 8 or comprises an amino acid sequence analogue having a sequence identity with any of said polypeptide sequences of at least 70% and at the same time being immunologically equivalent to said polypeptide sequence.

2. A substantially pure polypeptide fragment which comprises a T-cell epitope of the polypeptide sequence as shown in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 or SEQ ID NO: 8 and at the same time being immunologically equivalent to said polypeptide.

3. A fusion polypeptide comprising at least one polypeptide fragment according to any of the preceding claims and at least one fusion partner.

4. A fusion polypeptide according to claim 3, wherein the fusion partner is selected from the group consisting of a polypeptide fragment derived from a bacterium belonging to the tuberculosis complex, such as ESAT-6 or at least one T-cell epitope thereof, MPB64 or at least one T-cell epitope thereof, MPT64 or at least one T-cell epitope thereof, and MPB59 or at least one T-cell epitope thereof.

5. A polypeptide according to any of the preceding claims which is lipidated so as to allow a self-adjuvating effect of the polypeptide.

6. A substantially pure polypeptide according to any of the preceding clams for use as a pharmaceutical.

7. Use of a substantially pure polypeptide according to any of the preceding claims for the preparation of a pharmaceutical composition for the diagnosis of tuberculosis caused by *Mycobacterium tuberculosis*, *Mycobacterium africanum* or *Mycobacterium bovis*.

8. Use of a substantially pure polypeptide according to any of the preceding claims for the preparation of a pharmaceutical composition for the vaccination against infections caused by *Mycobacterium tuberculosis*, *Mycobacterium africanum* or *Mycobacterium bovis*.

9. An immunologic composition comprising a polypeptide according to any of the preceding claims.

10. An immunologic composition, which is in the form of a vaccine.

11. An immunologic composition, which is in the form of a skin test reagent.

12. A nucleic acid fragment in isolated form which

1) comprises a nucleic acid sequence which encodes a polypeptide as defined in any of claims 1-5, or comprises a nucleic acid sequence complementary thereto; or

2) has a length of at least 10 nucleotides and hybridizes readily under stringent hybridization conditions with a nucleic acid fragment which has a nucleotide sequence selected from
SEQ ID NO: 1 or a sequence complementary thereto,
SEQ ID NO: 3 or a sequence complementary thereto,
SEQ ID NO: 5 or a sequence complementary thereto, and
SEQ ID NO: 7 or a sequence complementary thereto.

13. A nucleic acid fragment according to claim 12, which is a DNA fragment.

14. A nucleic acid fragment according to claim 12 or 13 for use as a pharmaceutical.

15. A vaccine comprising a nucleic acid fragment according to claim 12 or 13, the vaccine effecting *in vivo* expression of antigen by an animal, including a human being, to whom the vaccine has been administered, the amount of expressed antigen being effective to confer substantially increased resistance to infections with mycobacteria of the tuberculosis complex in an animal, including a human being.

16. Use of a nucleic acid fragment according to claim 12 or 13 for the preparation of a pharmaceutical composition for the diagnosis of tuberculosis caused by *Mycobacterium tuberculosis*, *Mycobacterium africanum* or *Mycobacterium bovis*.

17. Use of a nucleic acid fragment according to claim 12 or 13 for the preparation of a pharmaceutical composition for the vaccination against tuberculosis caused by *Mycobacterium tuberculosis*, *Mycobacterium africanum* or *Mycobacterium bovis*.

18. A vaccine for immunizing an animal, including a human being, against tuberculosis caused by mycobacteria belonging to the tuberculosis complex, comprising as the effective component a non-pathogenic microorganism, wherein at least one copy of a DNA fragment comprising a DNA sequence encoding a polypeptide according to any of claims 1-5 has been incorporated into the genome of the microorganism in a manner allowing the microorganism to express and optionally secrete the polypeptide.

19. A replicable expression vector which comprises a nucleic acid fragment according to claim 12 or 13.

20. A transformed cell harbouring at least one vector according to claim 19.

21. A method for producing a polypeptide according to any of claims 1-5, comprising

inserting a nucleic acid fragment according to claim 12 or 13 into a vector which is able to replicate in a host cell, introducing the resulting recombinant vector into the host cell, culturing the host cell in a culture medium under conditions sufficient to effect expression of the polypeptide, and recovering the polypeptide from the host cell or culture medium; or

isolating the polypeptide from whole mycobacteria of the tuberculosis complex or from lysates or fractions thereof; or

synthesizing the polypeptide by solid or liquid phase peptide synthesis.

22. A method of diagnosing tuberculosis caused by *Mycobacterium tuberculosis*, *Mycobacterium africanum* or *Mycobacterium bovis* in an animal, including a human being, comprising intradermally injecting, in the animal, a polypeptide according to any of claims 1-5 or an immunologic composition according to claim 9, a positive skin response at the location of injection being indicative of the animal having tuberculosis, and a negative skin response at the location of injection being indicative of the animal not having tuberculosis.

23. A method for immunising an animal, including a human being, against tuberculosis caused by mycobacteria belonging to the tuberculosis complex, comprising administering to the animal the polypeptide according to any of claims 1-5, the immunologic composition according to claim 9, or the vaccine according to claim 18.

5

24. A monoclonal or polyclonal antibody, which is specifically reacting with a polypeptide according to any of claims 1-5 in an immuno assay, or a specific binding fragment of said antibody.

SEQUENCE LISTING

<110> Statens Serum Institut

<120> Nucleic acid fragments and polypeptide fragments derived from M. Tuberculosis

<130> 22030DK1

<160> 8

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 3990
<212> DNA
<213> Mycobacterium Tuberculosis

<220>
<221> CDS
<222> (1)...(3990)

<400> 1

```
gtg agc aga ctg atc ttt gag gct cgt cgc cga ctg gcg ccg ccg agc       48
Val Ser Arg Leu Ile Phe Glu Ala Arg Arg Arg Leu Ala Pro Pro Ser
 1               5                  10                  15

agc cac cag ggc acc atc atc atc gag gcg cct ccc gag ctg cct cgg       96
Ser His Gln Gly Thr Ile Ile Ile Glu Ala Pro Pro Glu Leu Pro Arg
            20                  25                  30

gtg atc cca ccg tca ctg cta cga cga gcg ctg cct tat ctg atc ggg      144
Val Ile Pro Pro Ser Leu Leu Arg Arg Ala Leu Pro Tyr Leu Ile Gly
        35                  40                  45

atc ctc atc gtg ggg atg atc gtg gcg ctg gtc gcc acc ggg atg cgg      192
Ile Leu Ile Val Gly Met Ile Val Ala Leu Val Ala Thr Gly Met Arg
    50                  55                  60

gtg att tct ccg cag acg ttg ttc ttc cca ttt gtg ctg ctg ttg gcg      240
Val Ile Ser Pro Gln Thr Leu Phe Phe Pro Phe Val Leu Leu Leu Ala
65                  70                  75                  80

gcc acc gcg ctc tac cgc ggc aac gac aag aag atg cgc acc gag gag      288
Ala Thr Ala Leu Tyr Arg Gly Asn Asp Lys Lys Met Arg Thr Glu Glu
                85                  90                  95

gtc gac gcc gaa cgg gcc gac tac cta cgt tac cta tcg gtg gtg cgg      336
Val Asp Ala Glu Arg Ala Asp Tyr Leu Arg Tyr Leu Ser Val Val Arg
            100                 105                 110

gac aac att cgg gcc cag gcc gcc gag cag cgg gcc agc gcg ttg tgg      384
Asp Asn Ile Arg Ala Gln Ala Ala Glu Gln Arg Ala Ser Ala Leu Trp
        115                 120                 125

tct cat cct gac ccg acg gcg ttg gcg tcg gtg ccg ggg tca cgt cgc      432
Ser His Pro Asp Pro Thr Ala Leu Ala Ser Val Pro Gly Ser Arg Arg
```

```
    130                 135                 140

caa tgg gag cgt gac ccg cac gac ccc gac ttt ttg gtg ttg cgg gcc      480
Gln Trp Glu Arg Asp Pro His Asp Pro Asp Phe Leu Val Leu Arg Ala
145                 150                 155                 160

ggc cgg cac acg gta ccg ctg gct act acg ctg cga gtc aac gac acc      528
Gly Arg His Thr Val Pro Leu Ala Thr Thr Leu Arg Val Asn Asp Thr
                165                 170                 175

gcc gac gag atc gac ctg gaa ccg gtg tcg cac agt gca tta cgc agc      576
Ala Asp Glu Ile Asp Leu Glu Pro Val Ser His Ser Ala Leu Arg Ser
            180                 185                 190

ctg ctc gac acc cag cgc agc att ggc gac gtg ccg acc ggg atc gac      624
Leu Leu Asp Thr Gln Arg Ser Ile Gly Asp Val Pro Thr Gly Ile Asp
        195                 200                 205

ctg acc aag gtt tcg ccg atc acc gtg ctg ggg gag cgc gca cag gtg      672
Leu Thr Lys Val Ser Pro Ile Thr Val Leu Gly Glu Arg Ala Gln Val
    210                 215                 220

cgc gcg gtg tta cgc gcc tgg atc gct cag gcg gtg acc tgg cac gac      720
Arg Ala Val Leu Arg Ala Trp Ile Ala Gln Ala Val Thr Trp His Asp
225                 230                 235                 240

ccg acg gtg ctc ggg gtg gcg ctg gcc gcg cgt gat ctg gag ggt cgc      768
Pro Thr Val Leu Gly Val Ala Leu Ala Ala Arg Asp Leu Glu Gly Arg
                245                 250                 255

gat tgg aac tgg ctg aag tgg tta ccg cac gtg gac att ccc ggc cgc      816
Asp Trp Asn Trp Leu Lys Trp Leu Pro His Val Asp Ile Pro Gly Arg
            260                 265                 270

ctc gat gcg ctg ggc ccg gcc cgc aat ctg tcg acc gat ccc gac gag      864
Leu Asp Ala Leu Gly Pro Ala Arg Asn Leu Ser Thr Asp Pro Asp Glu
        275                 280                 285

ctc atc gcg ctg ctg ggg ccc gtc ctg gca gac cgc ccg gcg ttt acc      912
Leu Ile Ala Leu Leu Gly Pro Val Leu Ala Asp Arg Pro Ala Phe Thr
    290                 295                 300

ggg cag cca aca gat gcg ttg cgg cac ttg ctg atc gtc gtc gat gac      960
Gly Gln Pro Thr Asp Ala Leu Arg His Leu Leu Ile Val Val Asp Asp
305                 310                 315                 320

ccg gac tac gac ctg ggc gca tcg ccg ctg gcg gtg ggc cgc gcg ggt     1008
Pro Asp Tyr Asp Leu Gly Ala Ser Pro Leu Ala Val Gly Arg Ala Gly
                325                 330                 335

gtc acc gtc gtg cac tgc tcg gcc agt gcg ccg cac cgg gaa cag tat     1056
Val Thr Val Val His Cys Ser Ala Ser Ala Pro His Arg Glu Gln Tyr
            340                 345                 350

tcg gat ccg gaa aag ccg atc ctg cgg gtg gct cac ggc gct atc gaa     1104
Ser Asp Pro Glu Lys Pro Ile Leu Arg Val Ala His Gly Ala Ile Glu
        355                 360                 365

cgc tgg cag aca ggc ggc tgg cag ccc tac atc gac gcc gcc gac caa     1152
```

```
Arg Trp Gln Thr Gly Gly Trp Gln Pro Tyr Ile Asp Ala Ala Asp Gln
    370                 375                 380

ttc agc gct gat gag gcc gcc cac ctg gcg cgc cga ctg tcg cgg tgg      1200
Phe Ser Ala Asp Glu Ala Ala His Leu Ala Arg Arg Leu Ser Arg Trp
385                 390                 395                 400

gac tcc aac ccc acc cat gcc ggg ctg cgc tcg gcg gcc act cgc ggc      1248
Asp Ser Asn Pro Thr His Ala Gly Leu Arg Ser Ala Ala Thr Arg Gly
                405                 410                 415

gcg agt ttc acc aca ctg ctg ggc atc gag gac gca tcc cga ctg gat      1296
Ala Ser Phe Thr Thr Leu Leu Gly Ile Glu Asp Ala Ser Arg Leu Asp
            420                 425                 430

gtg ccc gcg ctg tgg gcg ccg cga cga cgc gac gag gag tta cgc gtg      1344
Val Pro Ala Leu Trp Ala Pro Arg Arg Arg Asp Glu Glu Leu Arg Val
        435                 440                 445

ccg atc ggt gtc act ggc acc ggc gag ccg ctg atg ttc gac ctc aaa      1392
Pro Ile Gly Val Thr Gly Thr Gly Glu Pro Leu Met Phe Asp Leu Lys
    450                 455                 460

gac gaa gcc gag ggc ggg atg ggc ccg cac ggg ctg atg atc ggc atg      1440
Asp Glu Ala Glu Gly Gly Met Gly Pro His Gly Leu Met Ile Gly Met
465                 470                 475                 480

acc ggt tcg ggc aag tcg cag act ttg atg tcg att ctg ttg tcg ctg      1488
Thr Gly Ser Gly Lys Ser Gln Thr Leu Met Ser Ile Leu Leu Ser Leu
                485                 490                 495

ttg acc aca cac tcc gcg gag cgg ctc atc gtc atc tac gcc gac ttc      1536
Leu Thr Thr His Ser Ala Glu Arg Leu Ile Val Ile Tyr Ala Asp Phe
            500                 505                 510

aag ggt gag gcc ggc gcc gac agt ttc cga gat ttc ccg cag gtg gtt      1584
Lys Gly Glu Ala Gly Ala Asp Ser Phe Arg Asp Phe Pro Gln Val Val
        515                 520                 525

gcg gtg atc tcg aat atg gcc gag aag aag tcg ttg gct gat cgg ttc      1632
Ala Val Ile Ser Asn Met Ala Glu Lys Lys Ser Leu Ala Asp Arg Phe
    530                 535                 540

gcc gac acg ctg cgc ggc gag gtg gct cgt cgc gag atg ctg ctg cgt      1680
Ala Asp Thr Leu Arg Gly Glu Val Ala Arg Arg Glu Met Leu Leu Arg
545                 550                 555                 560

gag gcc ggc cgc aag gtc cag ggc agc gcg ttc aac tcg gtg ctc gag      1728
Glu Ala Gly Arg Lys Val Gln Gly Ser Ala Phe Asn Ser Val Leu Glu
                565                 570                 575

tat gaa aac gcc atc gcc gca ggg cat agc ctg ccg ccc atc ccg aca      1776
Tyr Glu Asn Ala Ile Ala Ala Gly His Ser Leu Pro Pro Ile Pro Thr
            580                 585                 590

ctg ttc gtg gtc gcc gac gag ttc acc ttg atg ctg gcc gat cac ccg      1824
Leu Phe Val Val Ala Asp Glu Phe Thr Leu Met Leu Ala Asp His Pro
        595                 600                 605
```

```
gaa tac gcg gag ctg ttc gac tat gtg gcc cgc aag ggt cgc tcg ttt     1872
Glu Tyr Ala Glu Leu Phe Asp Tyr Val Ala Arg Lys Gly Arg Ser Phe
    610                 615                 620

cgc atc cac atc cta ttc gcg tcc cag aca ctg gac gtg ggc aag atc     1920
Arg Ile His Ile Leu Phe Ala Ser Gln Thr Leu Asp Val Gly Lys Ile
625                 630                 635                 640

aaa gac atc gac aag aac acc gcc tat cgg att ggg ctg aaa gtg gcc     1968
Lys Asp Ile Asp Lys Asn Thr Ala Tyr Arg Ile Gly Leu Lys Val Ala
                645                 650                 655

agc ccc agc gtt tct cgc cag atc atc ggc gtg gag gac gcc tac cac     2016
Ser Pro Ser Val Ser Arg Gln Ile Ile Gly Val Glu Asp Ala Tyr His
            660                 665                 670

atc gag tcg ggc aaa gaa cac aaa ggc gtg ggc ttt ttg gtg ccc gcg     2064
Ile Glu Ser Gly Lys Glu His Lys Gly Val Gly Phe Leu Val Pro Ala
        675                 680                 685

ccc ggt gcc acc ccg ata agg ttc cgc agc acc tat gtc gac ggg atc     2112
Pro Gly Ala Thr Pro Ile Arg Phe Arg Ser Thr Tyr Val Asp Gly Ile
    690                 695                 700

tat gaa ccg ccg cag acg gct aaa gcc gtt gtc gtg caa tcc gtt ccg     2160
Tyr Glu Pro Pro Gln Thr Ala Lys Ala Val Val Val Gln Ser Val Pro
705                 710                 715                 720

gag ccc aag ctg ttc acc gcc gcc gcg gtg gaa ccg gat ccg ggc acg     2208
Glu Pro Lys Leu Phe Thr Ala Ala Ala Val Glu Pro Asp Pro Gly Thr
                725                 730                 735

gtg atc gcc gat act gac gaa caa gaa ccc gcc gac cca cca cgc aaa     2256
Val Ile Ala Asp Thr Asp Glu Gln Glu Pro Ala Asp Pro Pro Arg Lys
            740                 745                 750

ctg atc gcg acc atc ggc gaa caa ctg gcc cgc tac ggt ccg cgg gcg     2304
Leu Ile Ala Thr Ile Gly Glu Gln Leu Ala Arg Tyr Gly Pro Arg Ala
        755                 760                 765

ccg cag ttg tgg ctg ccg cca ctc gac gaa acg atc cca ctg agc gcg     2352
Pro Gln Leu Trp Leu Pro Pro Leu Asp Glu Thr Ile Pro Leu Ser Ala
    770                 775                 780

gcg ttg gcc cgc gcc ggg gtg ggc ccc cgg cag tgg cgc tgg ccg ctg     2400
Ala Leu Ala Arg Ala Gly Val Gly Pro Arg Gln Trp Arg Trp Pro Leu
785                 790                 795                 800

ggg gag atc gac agg ccc ttc gag atg cgg cgc gac ccg ttg gtg ttt     2448
Gly Glu Ile Asp Arg Pro Phe Glu Met Arg Arg Asp Pro Leu Val Phe
                805                 810                 815

gac gct agg tcg tcg gcc gga aat atg gtg atc cac ggc ggc ccc aag     2496
Asp Ala Arg Ser Ser Ala Gly Asn Met Val Ile His Gly Gly Pro Lys
            820                 825                 830

tcc ggc aaa tcc act gcg ctg cag aca ttc atc ctc tca gct gct agc     2544
Ser Gly Lys Ser Thr Ala Leu Gln Thr Phe Ile Leu Ser Ala Ala Ser
        835                 840                 845
```

```
ctg cac tcg ccg cac gag gtt agc ttc tat tgc ctg gac tac ggc ggt     2592
Leu His Ser Pro His Glu Val Ser Phe Tyr Cys Leu Asp Tyr Gly Gly
    850                 855                 860

ggg cag ctg cgg gcg cta cag gat cta gcg cac gtc ggc agt gtc gcc     2640
Gly Gln Leu Arg Ala Leu Gln Asp Leu Ala His Val Gly Ser Val Ala
865                 870                 875                 880

tca gcg ctg gaa ccc gaa cgc atc cgc cgc acc ttc ggc gag ctc gag     2688
Ser Ala Leu Glu Pro Glu Arg Ile Arg Arg Thr Phe Gly Glu Leu Glu
                885                 890                 895

caa ctg ctg ttg tcc cgg cag cag cgg gaa gta ttc cgt gac cgg ggt     2736
Gln Leu Leu Leu Ser Arg Gln Gln Arg Glu Val Phe Arg Asp Arg Gly
            900                 905                 910

gct aat ggc tcg acc ccc gac gac ggg ttc ggt gag gtg ttc ctg gtc     2784
Ala Asn Gly Ser Thr Pro Asp Asp Gly Phe Gly Glu Val Phe Leu Val
        915                 920                 925

atc gac aat ctc tat ggc ttc ggc cgc gat aac acc gat cag ttc aac     2832
Ile Asp Asn Leu Tyr Gly Phe Gly Arg Asp Asn Thr Asp Gln Phe Asn
    930                 935                 940

acc cgt aat ccg ttg ctg gcc agg gta acc gaa ctg gtc aac gtg ggc     2880
Thr Arg Asn Pro Leu Leu Ala Arg Val Thr Glu Leu Val Asn Val Gly
945                 950                 955                 960

ctt gcc tac ggg atc cac gtg atc att acc acg ccg agc tgg ctg gaa     2928
Leu Ala Tyr Gly Ile His Val Ile Ile Thr Thr Pro Ser Trp Leu Glu
                965                 970                 975

gtg ccg ttg gcg atg cgc gac ggg ctc ggg ctg cgt ctc gag ctg cga     2976
Val Pro Leu Ala Met Arg Asp Gly Leu Gly Leu Arg Leu Glu Leu Arg
            980                 985                 990

ctg cac gac gcg cgc gac agc aac gtg cgg gtg gtc ggc gcc ctg cgc     3024
Leu His Asp Ala Arg Asp Ser Asn Val Arg Val Val Gly Ala Leu Arg
        995                 1000                1005

cgc ccg gcc gac gcc gtc ccg cac gac cag ccc ggc cgc gga ctg acc     3072
Arg Pro Ala Asp Ala Val Pro His Asp Gln Pro Gly Arg Gly Leu Thr
    1010                1015                1020

atg gcc gcc gag cac ttc ctg ttc gcg gct cca gaa ctg gac gcg caa     3120
Met Ala Ala Glu His Phe Leu Phe Ala Ala Pro Glu Leu Asp Ala Gln
1025                1030                1035                1040

aca aac ccg gtg gcc gcg atc aac gcc cgc tac ccc ggc atg gcg gct     3168
Thr Asn Pro Val Ala Ala Ile Asn Ala Arg Tyr Pro Gly Met Ala Ala
                1045                1050                1055

ccc ccg gtt cgg ttg ttg ccc acc aac ctt gcg ccg cac gcc gtc ggc     3216
Pro Pro Val Arg Leu Leu Pro Thr Asn Leu Ala Pro His Ala Val Gly
            1060                1065                1070

gaa ctg tat cgg ggt ccc gac caa ctg gtg att ggc cag cgc gaa gaa     3264
Glu Leu Tyr Arg Gly Pro Asp Gln Leu Val Ile Gly Gln Arg Glu Glu
```

```
        1075                1080                1085

gac ctg gcg ccg gtg ata ctc gac ctc gcc gcc aac ccg ctg ctg atg      3312
Asp Leu Ala Pro Val Ile Leu Asp Leu Ala Ala Asn Pro Leu Leu Met
    1090                1095                1100

gtg ttc ggc gat gcc agg tca gga aag acg acg ctg ctg cgc cac atc      3360
Val Phe Gly Asp Ala Arg Ser Gly Lys Thr Thr Leu Leu Arg His Ile
1105                1110                1115                1120

atc cgc acc gtc cgc gag cac tcc acc gcc gac cgg gtc gcg ttc acc      3408
Ile Arg Thr Val Arg Glu His Ser Thr Ala Asp Arg Val Ala Phe Thr
                1125                1130                1135

gtg ctg gac cgc cgg cta cac ctg gtc gac gaa cca ctg ttc ccc gac      3456
Val Leu Asp Arg Arg Leu His Leu Val Asp Glu Pro Leu Phe Pro Asp
            1140                1145                1150

aac gag tac acc gcc aac atc gat cgg atc atc ccg gcg atg ctc ggg      3504
Asn Glu Tyr Thr Ala Asn Ile Asp Arg Ile Ile Pro Ala Met Leu Gly
        1155                1160                1165

ctg gcc aac ctc atc gag gcg cgc cgg ccg ccg gcc ggg atg tct gcg      3552
Leu Ala Asn Leu Ile Glu Ala Arg Arg Pro Pro Ala Gly Met Ser Ala
    1170                1175                1180

gcc gag ctg tcc cgc tgg acc ttt gcc ggg cac acc cac tac ctg atc      3600
Ala Glu Leu Ser Arg Trp Thr Phe Ala Gly His Thr His Tyr Leu Ile
1185                1190                1195                1200

atc gac gac gtc gac cag gta ccg gat tcg ccg gcg atg acc ggt ccc      3648
Ile Asp Asp Val Asp Gln Val Pro Asp Ser Pro Ala Met Thr Gly Pro
                1205                1210                1215

tac atc gga cag cgg ccg tgg acc ccg ctg atc ggt ctc ctg gcc cag      3696
Tyr Ile Gly Gln Arg Pro Trp Thr Pro Leu Ile Gly Leu Leu Ala Gln
            1220                1225                1230

gcc ggc gac ttg ggg cta cgg gtg att gtc acc ggg cgt gcc act gga      3744
Ala Gly Asp Leu Gly Leu Arg Val Ile Val Thr Gly Arg Ala Thr Gly
        1235                1240                1245

tcg gcg cac ctg ctg atg aca agt ccg ttg ctg cgc cgg ttc aac gac      3792
Ser Ala His Leu Leu Met Thr Ser Pro Leu Leu Arg Arg Phe Asn Asp
    1250                1255                1260

ctg cag gcg acc acg ctg atg ttg gca ggc aat ccg gcc gac agc ggc      3840
Leu Gln Ala Thr Thr Leu Met Leu Ala Gly Asn Pro Ala Asp Ser Gly
1265                1270                1275                1280

aag att cgc ggt gag cgg ttt gcc cga ttg cct gct gga cga gca att      3888
Lys Ile Arg Gly Glu Arg Phe Ala Arg Leu Pro Ala Gly Arg Ala Ile
                1285                1290                1295

ctg ttg acc gac agt gat agt cca acc tac gtg cag ttg atc aac ccg      3936
Leu Leu Thr Asp Ser Asp Ser Pro Thr Tyr Val Gln Leu Ile Asn Pro
            1300                1305                1310

ctg gtc gat gcg gcc gcg gtt tct ggt gaa acc caa cag aag ggg agt      3984
```

Leu Val Asp Ala Ala Ala Val Ser Gly Glu Thr Gln Gln Lys Gly Ser
1315 1320 1325

cag tca 3990
Gln Ser
1330

<210> 2
<211> 1330
<212> PRT
<213> Mycobacterium Tuberculosis

<400> 2

```
Val Ser Arg Leu Ile Phe Glu Ala Arg Arg Arg Leu Ala Pro Pro Ser
 1               5                  10                  15
Ser His Gln Gly Thr Ile Ile Ile Glu Ala Pro Pro Glu Leu Pro Arg
            20                  25                  30
Val Ile Pro Pro Ser Leu Leu Arg Arg Ala Leu Pro Tyr Leu Ile Gly
        35                  40                  45
Ile Leu Ile Val Gly Met Ile Val Ala Leu Val Ala Thr Gly Met Arg
    50                  55                  60
Val Ile Ser Pro Gln Thr Leu Phe Phe Pro Phe Val Leu Leu Leu Ala
65                  70                  75                  80
Ala Thr Ala Leu Tyr Arg Gly Asn Asp Lys Lys Met Arg Thr Glu Glu
                85                  90                  95
Val Asp Ala Glu Arg Ala Asp Tyr Leu Arg Tyr Leu Ser Val Val Arg
            100                 105                 110
Asp Asn Ile Arg Ala Gln Ala Ala Glu Gln Arg Ala Ser Ala Leu Trp
        115                 120                 125
Ser His Pro Asp Pro Thr Ala Leu Ala Ser Val Pro Gly Ser Arg Arg
    130                 135                 140
Gln Trp Glu Arg Asp Pro His Asp Pro Asp Phe Leu Val Leu Arg Ala
145                 150                 155                 160
Gly Arg His Thr Val Pro Leu Ala Thr Thr Leu Arg Val Asn Asp Thr
                165                 170                 175
Ala Asp Glu Ile Asp Leu Glu Pro Val Ser His Ser Ala Leu Arg Ser
            180                 185                 190
Leu Leu Asp Thr Gln Arg Ser Ile Gly Asp Val Pro Thr Gly Ile Asp
        195                 200                 205
Leu Thr Lys Val Ser Pro Ile Thr Val Leu Gly Glu Arg Ala Gln Val
    210                 215                 220
Arg Ala Val Leu Arg Ala Trp Ile Ala Gln Ala Val Thr Trp His Asp
225                 230                 235                 240
Pro Thr Val Leu Gly Val Ala Leu Ala Ala Arg Asp Leu Glu Gly Arg
                245                 250                 255
Asp Trp Asn Trp Leu Lys Trp Leu Pro His Val Asp Ile Pro Gly Arg
            260                 265                 270
Leu Asp Ala Leu Gly Pro Ala Arg Asn Leu Ser Thr Asp Pro Asp Glu
        275                 280                 285
Leu Ile Ala Leu Leu Gly Pro Val Leu Ala Asp Arg Pro Ala Phe Thr
    290                 295                 300
Gly Gln Pro Thr Asp Ala Leu Arg His Leu Leu Ile Val Val Asp Asp
305                 310                 315                 320
Pro Asp Tyr Asp Leu Gly Ala Ser Pro Leu Ala Val Gly Arg Ala Gly
                325                 330                 335
Val Thr Val Val His Cys Ser Ala Ser Ala Pro His Arg Glu Gln Tyr
            340                 345                 350
Ser Asp Pro Glu Lys Pro Ile Leu Arg Val Ala His Gly Ala Ile Glu
```

```
        355                 360                 365
Arg Trp Gln Thr Gly Gly Trp Gln Pro Tyr Ile Asp Ala Ala Asp Gln
    370                 375                 380
Phe Ser Ala Asp Glu Ala Ala His Leu Ala Arg Arg Leu Ser Arg Trp
385                 390                 395                 400
Asp Ser Asn Pro Thr His Ala Gly Leu Arg Ser Ala Ala Thr Arg Gly
                405                 410                 415
Ala Ser Phe Thr Thr Leu Leu Gly Ile Glu Asp Ala Ser Arg Leu Asp
            420                 425                 430
Val Pro Ala Leu Trp Ala Pro Arg Arg Arg Asp Glu Glu Leu Arg Val
        435                 440                 445
Pro Ile Gly Val Thr Gly Thr Gly Glu Pro Leu Met Phe Asp Leu Lys
    450                 455                 460
Asp Glu Ala Glu Gly Gly Met Gly Pro His Gly Leu Met Ile Gly Met
465                 470                 475                 480
Thr Gly Ser Gly Lys Ser Gln Thr Leu Met Ser Ile Leu Leu Ser Leu
                485                 490                 495
Leu Thr Thr His Ser Ala Glu Arg Leu Ile Val Ile Tyr Ala Asp Phe
            500                 505                 510
Lys Gly Glu Ala Gly Ala Asp Ser Phe Arg Asp Phe Pro Gln Val Val
        515                 520                 525
Ala Val Ile Ser Asn Met Ala Glu Lys Lys Ser Leu Ala Asp Arg Phe
    530                 535                 540
Ala Asp Thr Leu Arg Gly Glu Val Ala Arg Arg Glu Met Leu Leu Arg
545                 550                 555                 560
Glu Ala Gly Arg Lys Val Gln Gly Ser Ala Phe Asn Ser Val Leu Glu
                565                 570                 575
Tyr Glu Asn Ala Ile Ala Ala Gly His Ser Leu Pro Pro Ile Pro Thr
            580                 585                 590
Leu Phe Val Val Ala Asp Glu Phe Thr Leu Met Leu Ala Asp His Pro
        595                 600                 605
Glu Tyr Ala Glu Leu Phe Asp Tyr Val Ala Arg Lys Gly Arg Ser Phe
    610                 615                 620
Arg Ile His Ile Leu Phe Ala Ser Gln Thr Leu Asp Val Gly Lys Ile
625                 630                 635                 640
Lys Asp Ile Asp Lys Asn Thr Ala Tyr Arg Ile Gly Leu Lys Val Ala
                645                 650                 655
Ser Pro Ser Val Ser Arg Gln Ile Ile Gly Val Glu Asp Ala Tyr His
            660                 665                 670
Ile Glu Ser Gly Lys Glu His Lys Gly Val Gly Phe Leu Val Pro Ala
        675                 680                 685
Pro Gly Ala Thr Pro Ile Arg Phe Arg Ser Thr Tyr Val Asp Gly Ile
    690                 695                 700
Tyr Glu Pro Pro Gln Thr Ala Lys Ala Val Val Val Gln Ser Val Pro
705                 710                 715                 720
Glu Pro Lys Leu Phe Thr Ala Ala Ala Val Glu Pro Asp Pro Gly Thr
                725                 730                 735
Val Ile Ala Asp Thr Asp Glu Gln Glu Pro Ala Asp Pro Pro Arg Lys
            740                 745                 750
Leu Ile Ala Thr Ile Gly Glu Gln Leu Ala Arg Tyr Gly Pro Arg Ala
        755                 760                 765
Pro Gln Leu Trp Leu Pro Pro Leu Asp Glu Thr Ile Pro Leu Ser Ala
    770                 775                 780
Ala Leu Ala Arg Ala Gly Val Gly Pro Arg Gln Trp Arg Trp Pro Leu
785                 790                 795                 800
Gly Glu Ile Asp Arg Pro Phe Glu Met Arg Arg Asp Pro Leu Val Phe
                805                 810                 815
Asp Ala Arg Ser Ser Ala Gly Asn Met Val Ile His Gly Gly Pro Lys
            820                 825                 830
```

```
Ser Gly Lys Ser Thr Ala Leu Gln Thr Phe Ile Leu Ser Ala Ala Ser
        835                 840                 845
Leu His Ser Pro His Glu Val Ser Phe Tyr Cys Leu Asp Tyr Gly Gly
    850                 855                 860
Gly Gln Leu Arg Ala Leu Gln Asp Leu Ala His Val Gly Ser Val Ala
865                 870                 875                 880
Ser Ala Leu Glu Pro Glu Arg Ile Arg Arg Thr Phe Gly Glu Leu Glu
                885                 890                 895
Gln Leu Leu Leu Ser Arg Gln Gln Arg Glu Val Phe Arg Asp Arg Gly
            900                 905                 910
Ala Asn Gly Ser Thr Pro Asp Asp Gly Phe Gly Glu Val Phe Leu Val
        915                 920                 925
Ile Asp Asn Leu Tyr Gly Phe Gly Arg Asp Asn Thr Asp Gln Phe Asn
    930                 935                 940
Thr Arg Asn Pro Leu Leu Ala Arg Val Thr Glu Leu Val Asn Val Gly
945                 950                 955                 960
Leu Ala Tyr Gly Ile His Val Ile Ile Thr Thr Pro Ser Trp Leu Glu
                965                 970                 975
Val Pro Leu Ala Met Arg Asp Gly Leu Gly Leu Arg Leu Glu Leu Arg
            980                 985                 990
Leu His Asp Ala Arg Asp Ser Asn Val Arg Val Val Gly Ala Leu Arg
        995                 1000                1005
Arg Pro Ala Asp Ala Val Pro His Asp Gln Pro Gly Arg Gly Leu Thr
    1010                1015                1020
Met Ala Ala Glu His Phe Leu Phe Ala Ala Pro Glu Leu Asp Ala Gln
1025                1030                1035                1040
Thr Asn Pro Val Ala Ala Ile Asn Ala Arg Tyr Pro Gly Met Ala Ala
                1045                1050                1055
Pro Pro Val Arg Leu Leu Pro Thr Asn Leu Ala Pro His Ala Val Gly
            1060                1065                1070
Glu Leu Tyr Arg Gly Pro Asp Gln Leu Val Ile Gly Gln Arg Glu Glu
        1075                1080                1085
Asp Leu Ala Pro Val Ile Leu Asp Leu Ala Ala Asn Pro Leu Leu Met
    1090                1095                1100
Val Phe Gly Asp Ala Arg Ser Gly Lys Thr Thr Leu Leu Arg His Ile
1105                1110                1115                1120
Ile Arg Thr Val Arg Glu His Ser Thr Ala Asp Arg Val Ala Phe Thr
                1125                1130                1135
Val Leu Asp Arg Arg Leu His Leu Val Asp Glu Pro Leu Phe Pro Asp
            1140                1145                1150
Asn Glu Tyr Thr Ala Asn Ile Asp Arg Ile Ile Pro Ala Met Leu Gly
        1155                1160                1165
Leu Ala Asn Leu Ile Glu Ala Arg Arg Pro Pro Ala Gly Met Ser Ala
    1170                1175                1180
Ala Glu Leu Ser Arg Trp Thr Phe Ala Gly His Thr His Tyr Leu Ile
1185                1190                1195                1200
Ile Asp Asp Val Asp Gln Val Pro Asp Ser Pro Ala Met Thr Gly Pro
                1205                1210                1215
Tyr Ile Gly Gln Arg Pro Trp Thr Pro Leu Ile Gly Leu Leu Ala Gln
            1220                1225                1230
Ala Gly Asp Leu Gly Leu Arg Val Ile Val Thr Gly Arg Ala Thr Gly
        1235                1240                1245
Ser Ala His Leu Leu Met Thr Ser Pro Leu Leu Arg Arg Phe Asn Asp
    1250                1255                1260
Leu Gln Ala Thr Thr Leu Met Leu Ala Gly Asn Pro Ala Asp Ser Gly
1265                1270                1275                1280
Lys Ile Arg Gly Glu Arg Phe Ala Arg Leu Pro Ala Gly Arg Ala Ile
                1285                1290                1295
Leu Leu Thr Asp Ser Asp Ser Pro Thr Tyr Val Gln Leu Ile Asn Pro
```

```
          1300                1305                1310
Leu Val Asp Ala Ala Ala Val Ser Gly Glu Thr Gln Gln Lys Gly Ser
       1315                1320                1325
Gln Ser
    1330
```

<210> 3
<211> 375
<212> DNA
<213> Mycobacterium Tuberculosis

<220>
<221> CDS
<222> (1)...(375)

<400> 3

```
cag gta ccg gat tcg ccg gcg atg acc ggt ccc tac atc gga cag cgg        48
Gln Val Pro Asp Ser Pro Ala Met Thr Gly Pro Tyr Ile Gly Gln Arg
 1               5                  10                  15

ccg tgg acc ccg ctg atc ggt ctc ctg gcc cag gcc ggc gac ttg ggg        96
Pro Trp Thr Pro Leu Ile Gly Leu Leu Ala Gln Ala Gly Asp Leu Gly
            20                  25                  30

cta cgg gtg att gtc acc ggg cgt gcc act gga tcg gcg cac ctg ctg       144
Leu Arg Val Ile Val Thr Gly Arg Ala Thr Gly Ser Ala His Leu Leu
        35                  40                  45

atg aca agt ccg ttg ctg cgc cgg ttc aac gac ctg cag gcg acc acg       192
Met Thr Ser Pro Leu Leu Arg Arg Phe Asn Asp Leu Gln Ala Thr Thr
    50                  55                  60

ctg atg ttg gca ggc aat ccg gcc gac agc ggc aag att cgc ggt gag       240
Leu Met Leu Ala Gly Asn Pro Ala Asp Ser Gly Lys Ile Arg Gly Glu
65                  70                  75                  80

cgg ttt gcc cga ttg cct gct gga cga gca att ctg ttg acc gac agt       288
Arg Phe Ala Arg Leu Pro Ala Gly Arg Ala Ile Leu Leu Thr Asp Ser
                85                  90                  95

gat agt cca acc tac gtg cag ttg atc aac ccg ctg gtc gat gcg gcc       336
Asp Ser Pro Thr Tyr Val Gln Leu Ile Asn Pro Leu Val Asp Ala Ala
            100                 105                 110

gcg gtt tct ggt gaa acc caa cag aag ggg agt cag tca                  375
Ala Val Ser Gly Glu Thr Gln Gln Lys Gly Ser Gln Ser
        115                 120                 125
```

<210> 4
<211> 125
<212> PRT
<213> Mycobacterium Tuberculosis

<400> 4

```
Gln Val Pro Asp Ser Pro Ala Met Thr Gly Pro Tyr Ile Gly Gln Arg
 1               5                  10                  15
Pro Trp Thr Pro Leu Ile Gly Leu Leu Ala Gln Ala Gly Asp Leu Gly
            20                  25                  30
```

```
Leu Arg Val Ile Val Thr Gly Arg Ala Thr Gly Ser Ala His Leu Leu
      35                  40                  45
Met Thr Ser Pro Leu Leu Arg Arg Phe Asn Asp Leu Gln Ala Thr Thr
    50                  55                  60
Leu Met Leu Ala Gly Asn Pro Ala Asp Ser Gly Lys Ile Arg Gly Glu
65                  70                  75                  80
Arg Phe Ala Arg Leu Pro Ala Gly Arg Ala Ile Leu Leu Thr Asp Ser
                85                  90                  95
Asp Ser Pro Thr Tyr Val Gln Leu Ile Asn Pro Leu Val Asp Ala Ala
            100                 105                 110
Ala Val Ser Gly Glu Thr Gln Gln Lys Gly Ser Gln Ser
        115                 120                 125
```

<210> 5
<211> 306
<212> DNA
<213> Mycobacterium Tuberculosis

<220>
<221> CDS
<222> (1)...(306)

<400> 5

```
atg acg ttg cga gtg gtt ccg gag ggg ctg gcc gca gcc agc gct gcg      48
Met Thr Leu Arg Val Val Pro Glu Gly Leu Ala Ala Ala Ser Ala Ala
 1               5                  10                  15

gtg gaa gcg ctg acg gcg cgg ttg gcc gcc gcg cat gcg agc gca gcg      96
Val Glu Ala Leu Thr Ala Arg Leu Ala Ala Ala His Ala Ser Ala Ala
            20                  25                  30

ccg gtg att acc gcg gta gtg ccg ccg gcg gcg gat ccg gtg tcg ctg     144
Pro Val Ile Thr Ala Val Val Pro Pro Ala Ala Asp Pro Val Ser Leu
        35                  40                  45

cag acc gcg gcc ggg ttc agt gca cag ggc gtc gag cac gcg gtc gtc     192
Gln Thr Ala Ala Gly Phe Ser Ala Gln Gly Val Glu His Ala Val Val
    50                  55                  60

acc gcc gaa ggt gtc gaa gag ctg gga cgc gcc ggc gtt ggt gtg ggc     240
Thr Ala Glu Gly Val Glu Glu Leu Gly Arg Ala Gly Val Gly Val Gly
65                  70                  75                  80

gaa tcc ggc gcc agc tac ctg gcc ggt gat gcg gcc gcc gcc gct acg     288
Glu Ser Gly Ala Ser Tyr Leu Ala Gly Asp Ala Ala Ala Ala Ala Thr
                85                  90                  95

tac ggg gtc gtg ggc ggc                                            306
Tyr Gly Val Val Gly Gly
            100
```

<210> 6
<211> 102
<212> PRT
<213> Mycobacterium Tuberculosis

<400> 6

```
Met Thr Leu Arg Val Val Pro Glu Gly Leu Ala Ala Ala Ser Ala Ala
```

```
 1               5                  10                  15
Val Glu Ala Leu Thr Ala Arg Leu Ala Ala Ala His Ala Ser Ala Ala
            20                  25                  30
Pro Val Ile Thr Ala Val Val Pro Pro Ala Ala Asp Pro Val Ser Leu
        35                  40                  45
Gln Thr Ala Ala Gly Phe Ser Ala Gln Gly Val Glu His Ala Val Val
    50                  55                  60
Thr Ala Glu Gly Val Glu Glu Leu Gly Arg Ala Gly Val Gly Val Gly
65                  70                  75                  80
Glu Ser Gly Ala Ser Tyr Leu Ala Gly Asp Ala Ala Ala Ala Ala Thr
                85                  90                  95
Tyr Gly Val Val Gly Gly
            100
```

<210> 7
<211> 840
<212> DNA
<213> Mycobacterium Tuberculosis

<220>
<221> CDS
<222> (1)...(840)

<400> 7

```
atg gct gaa ccg ttg gcc gtc gat ccc acc ggc ttg agc gca gcg gcc      48
Met Ala Glu Pro Leu Ala Val Asp Pro Thr Gly Leu Ser Ala Ala Ala
 1               5                  10                  15

gcg aaa ttg gcc ggc ctc gtt ttt ccg cag cct ccg gcg ccg atc gcg      96
Ala Lys Leu Ala Gly Leu Val Phe Pro Gln Pro Pro Ala Pro Ile Ala
            20                  25                  30

gtc agc gga acg gat tcg gtg gta gca gca atc aac gag acc atg cca     144
Val Ser Gly Thr Asp Ser Val Val Ala Ala Ile Asn Glu Thr Met Pro
        35                  40                  45

agc atc gaa tcg ctg gtc agt gac ggg ctg ccc ggc gtg aaa gcc gcc     192
Ser Ile Glu Ser Leu Val Ser Asp Gly Leu Pro Gly Val Lys Ala Ala
    50                  55                  60

ctg act cga aca gca tcc aac atg aac gcg gcg gcg gac gtc tat gcg     240
Leu Thr Arg Thr Ala Ser Asn Met Asn Ala Ala Ala Asp Val Tyr Ala
65                  70                  75                  80

aag acc gat cag tca ctg gga acc agt ttg agc cag tat gca ttc ggc     288
Lys Thr Asp Gln Ser Leu Gly Thr Ser Leu Ser Gln Tyr Ala Phe Gly
                85                  90                  95

tcg tcg ggc gaa ggc ctg gct ggc gtc gcc tcg gtc ggt ggt cag cca     336
Ser Ser Gly Glu Gly Leu Ala Gly Val Ala Ser Val Gly Gly Gln Pro
            100                 105                 110

agt cag gct acc cag ctg ctg agc aca ccc gtg tca cag gtc acg acc     384
Ser Gln Ala Thr Gln Leu Leu Ser Thr Pro Val Ser Gln Val Thr Thr
        115                 120                 125

cag ctc ggc gag acg gcc gct gag ctg gca ccc cgt gtt gtt gcg acg     432
Gln Leu Gly Glu Thr Ala Ala Glu Leu Ala Pro Arg Val Val Ala Thr
    130                 135                 140
```

```
        115                 120                 125
Gln Leu Gly Glu Thr Ala Ala Glu Leu Ala Pro Arg Val Val Ala Thr
    130                 135                 140
Val Pro Gln Leu Val Gln Leu Ala Pro His Ala Val Gln Met Ser Gln
145                 150                 155                 160
Asn Ala Ser Pro Ile Ala Gln Thr Ile Ser Gln Thr Ala Gln Gln Ala
                165                 170                 175
Ala Gln Ser Ala Gln Gly Gly Ser Gly Pro Met Pro Ala Gln Leu Ala
            180                 185                 190
Ser Ala Glu Lys Pro Ala Thr Glu Gln Ala Glu Pro Val His Glu Val
        195                 200                 205
Thr Asn Asp Asp Gln Gly Asp Gln Gly Asp Val Gln Pro Ala Glu Val
    210                 215                 220
Val Ala Ala Ala Arg Asp Glu Gly Ala Gly Ala Ser Pro Gly Gln Gln
225                 230                 235                 240
Pro Gly Gly Gly Val Pro Ala Gln Ala Met Asp Thr Gly Ala Gly Ala
                245                 250                 255
Arg Pro Ala Ala Ser Pro Leu Ala Ala Pro Val Asp Pro Ser Thr Pro
            260                 265                 270
Ala Pro Ser Thr Thr Thr Thr Leu
        275                 280
```